

2015

Development of the Exams Data Analysis Spreadsheet as a Tool To Help Instructors Conduct Customizable Analyses of Student ACS Exam Data


Alexandra Brandriet

Iowa State University

Thomas Holme

Iowa State University, taholme@iastate.edu

Follow this and additional works at: http://lib.dr.iastate.edu/chem_pubs

 Part of the [Other Chemistry Commons](#), and the [Science and Mathematics Education Commons](#)

The complete bibliographic information for this item can be found at http://lib.dr.iastate.edu/chem_pubs/381. For information on how to cite this item, please visit <http://lib.dr.iastate.edu/howtocite.html>.

This Article is brought to you for free and open access by the Chemistry at Iowa State University Digital Repository. It has been accepted for inclusion in Chemistry Publications by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

Development of the Exams Data Analysis Spreadsheet as a Tool To Help Instructors Conduct Customizable Analyses of Student ACS Exam Data

Abstract

The American Chemical Society Examinations Institute (ACS-EI) has recently developed the Exams Data Analysis Spread (EDAS) as a tool to help instructors conduct customizable analyses of their student data from ACS exams. The EDAS calculations allow instructors to analyze their students' performances both at the total score and individual item levels, while also providing national normative results that can be used for comparison. Additionally, instructors can analyze results based on subsets of items of their choosing or items based on the "big ideas" from the Anchoring Concepts Content Map (ACCM). In order to evaluate the utility and usability of the EDAS for instructors, the EDAS went through trial testing with 10 chemistry instructors from across the country. The instructor feedback confirmed that the EDAS has multiple implications for the classroom and departmental assessment, but some additional revisions were needed to increase its usability. This feedback was also used to make a video user-guide that will help instructors through specific difficulties described during trial testing. Currently, an EDAS tool has been developed for the GC12F, GC10S, and GC13 exams.

Disciplines

Other Chemistry | Science and Mathematics Education

Comments

Reprinted (adapted) with permission from J. Chem. Educ., 2015, 92 (12), pp 2054–2061. Copyright 2015 American Chemical Society.

Development of the Exams Data Analysis Spreadsheet as a Tool To Help Instructors Conduct Customizable Analyses of Student ACS Exam Data

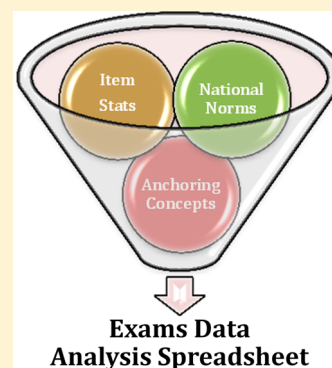
Alexandra Brandriet and Thomas Holme*

Department of Chemistry, Iowa State University, Ames, Iowa 50011, United States

S Supporting Information

ABSTRACT: The American Chemical Society Examinations Institute (ACS-EI) has recently developed the Exams Data Analysis Spread (EDAS) as a tool to help instructors conduct customizable analyses of their student data from ACS exams. The EDAS calculations allow instructors to analyze their students' performances both at the total score and individual item levels, while also providing national normative results that can be used for comparison. Additionally, instructors can analyze results based on subsets of items of their choosing or items based on the "big ideas" from the Anchoring Concepts Content Map (ACCM). In order to evaluate the utility and usability of the EDAS for instructors, the EDAS went through trial testing with 10 chemistry instructors from across the country. The instructor feedback confirmed that the EDAS has multiple implications for the classroom and departmental assessment, but some additional revisions were needed to increase its usability. This feedback was also used to make a video user-guide that will help instructors through specific difficulties described during trial testing. Currently, an EDAS tool has been developed for the GC12F, GC10S, and GC13 exams.

KEYWORDS: General Public, Computer-Based Learning, Testing/Assessment



ASSESSMENT IN CHEMISTRY

With movements to increase the quality of chemistry programs and aid in professional growth,¹ assessment plans and data-driven reform have become especially important topics of discussion within chemistry departments.^{2,3} As a result, more and more faculty are expected to become involved in programmatic assessment.² Despite the increasing demand, many have reported assessment-related challenges. In an editorial, Pienta¹ recounts that "A few too many people seemed to be saying that they didn't know much about assessment, how to do it, or even why they should." Certainly these sentiments are understood, since developing and implementing programmatic assessment can be a complex process that may involve multiple steps; Towns² describes a cyclic model of assessment that includes:

- (1) Setting the goals and priorities of the assessment effort.
- (2) Translating the goals into objectives.
- (3) Designing and implementing the assessment.
- (4) Evaluating the assessment findings.
- (5) Using the results to make decisions.

In an editorial, Bretz³ reiterates these challenges by stating that faculty may feel that they "...need a translator who is fluent in both chemistry and assessment." Such challenges may leave many faculty who are interested in data-driven reform wondering, "What do I do next?"

Despite the difficulties reported in the literature, assessment efforts have become quite common within chemistry departments. The American Chemical Society Examinations Institute

(ACS-EI) administered a national survey related to the current state of assessment practices and instructor needs within chemistry departments.^{4–8} On the basis of the responses from chemistry instructors from across the country ($N = 1,546$), Emenike and colleagues⁵ found that 72% reported that their department was expected to enhance assessment efforts or prepare reports related to assessment. Using cluster analysis, Raker and Holme⁸ identified distinct groups of instructors with differing levels of familiarity with assessment terminology. In this study, the instructor group with the highest degree of self-report familiarity only composed 20% of the sample. The largest group of instructors (25%) reported only a moderate understanding of assessment terminology, and specifically, a low level of familiarity with statistical analysis strategies.⁸ These results suggest that despite the commonality of departmental assessment, finding time to gain expertise in the evaluation of assessment data may be a hurdle for some faculty involved in these efforts.

TRANSLATING RESEARCH INTO PRACTICE

Evidence from survey research from the ACS-EI suggest that for some chemistry instructors, assessment-related jargon and methods may lead them to perceive it as an unfamiliar endeavor. As a result of such challenges, translating research into practice has become more important than ever. The National Research Council's report on Discipline-Based

Published: September 30, 2015

Education Research (DBER) describes that “Strategies are needed to more effectively promote the translation of findings from DBER into practice (Conclusion 12, p. 194),”⁹ and efforts to do this have become a common trend for the ACS-EI for several years. In response to instructor interest in the analysis of students’ content knowledge, the Anchoring Concepts Content Map (ACCM) was developed as a tool for programmatic assessment using ACS exams.^{10–14} More recently the ACS-EI has developed the Exams Data Analysis Spreadsheet (EDAS), which is a tool that instructors can use to help analyze their students’ ACS exam data. The EDAS features were developed with instructors in mind, based on both feedback from ACS-EI workshops and the results from the national survey of instructors’ assessment needs.^{4–8} Graphs a–c in Figure 1 show the responses of 1,067 instructors who responded to items related to their use of ACS exams.

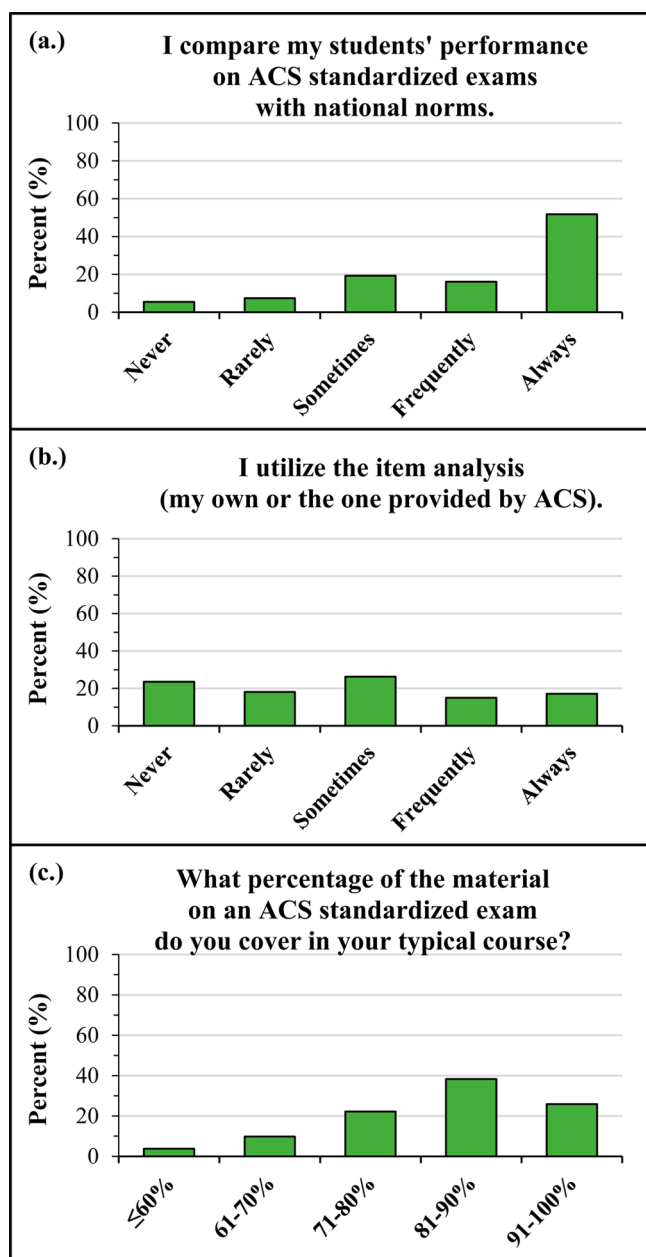


Figure 1. Instructors’ responses to items from the national survey of assessment needs.

The data analyses features within the EDAS were developed to help instructors continue to conduct the analyses that they have traditionally valued, such as comparing students’ performances to national norms (Figure 1a) while also giving them a tool to perform additional analyses that are apparently less common, such as examining individual item results (Figure 1b). Historically, national comparisons have highlighted overall test performances (i.e., total scores); however, the increasing demand for departmental assessment emphasizes the need to go beyond the whole-test level. The recent release of the 2015 guidelines for ACS program approval by the Committee on Professional Training^{15,16} have included important content recommendations, and ACS exams are excellent tools for assessing these. However, content-level guidelines may be better evaluated based on subsets of individual items, and the goal for the development of the EDAS was to simplify these types of analyses for instructors. As a result, the EDAS allows instructors to choose specific subsets of items or “big ideas” (i.e., anchoring concepts) based on the ACCM.^{10–14} Analogously, instructors can also choose to remove from their analyses specific items on topics that they do not teach in their courses. Because the ACS-EI acknowledges that individual classroom content may vary somewhat from ACS exam content (Figure 1c), instructors can choose to remove specific items from analyses and explore the differences this makes to their assessment results.

■ DEVELOPMENT OF THE EDAS

Features and Data Analyses

The EDAS was developed with user-friendliness as a primary goal. Because Microsoft (MS) Excel¹⁷ is a widely used software and many instructors likely have some familiarity with it, it was used as the platform to develop the EDAS. The data analyses are calculated using Excel macros, which are coded using Visual Basic for Applications (VBA) programming.¹⁸ The user does not need to interact with the coding; instead, the EDAS interface has multiple clickable buttons, radio-buttons, and check-boxes that instructors can use to interact with the data and customize analyses. This allows instructors to choose subsets of questions based either on their own interests or the anchoring concepts from the ACCM.^{10–14} Figure 2 shows a visual of the radio-buttons that can be used to toggle back-and-forth between individual items or anchoring concepts. Instructors can choose the specific items or anchoring concepts, as shown in Figures 3 and 4, by clicking or unclicking check-boxes. After the items or anchoring concepts are chosen, the user may then click a single button, as shown in Figure 2, and multiple results become available. A progress bar (Figure 2) helps the user know when the calculations are finished by progressing from “Not Started” to “Running” to “Finished.”

The EDAS has integrated several data analysis features at both the total score and individual item levels. At the total score level, the EDAS can calculate descriptive sample statistics (i.e., average, standard deviation, median, minimum, and maximum), frequencies of total scores, histograms, and percentiles. At the item level, difficulty and discrimination indices, as well as the percentage of students’ A, B, C, and D responses to each question are calculated. The customizable analyses can be conducted for both the instructor entered student data set and the national data set, and as a result, comparisons can be made across students’ performances and the national norms. A screenshot of some of the results provided by the EDAS are

Macro Progress:

Not Started

Calculate percentiles based on either:

☒ Individual Questions
 ☐ Anchoring Concepts
(choose from below)

Analyze Results

Figure 2. Radio-buttons used to customize analyses and the data analysis progress bar on the EDAS.

shown in Figure 5. The accuracy of the EDAS calculations were assessed by running multiple scenarios using SAS 9.4¹⁹ that encompassed various combinations of exam questions and anchoring concepts. Equivalent results were obtained from the EDAS as those provided by SAS.

Anchoring Concepts

☒ I. Atoms
☒ II. Bonding
☒ III. Structure and Function
☒ IV. Intermolecular Interactions
☒ V. Chemical Reactions
☒ VI. Energy and Thermodynamics
☐ VII. Kinetics
☐ VIII. Equilibrium
☒ IX. Experiments, Measurements, and Data
☒ X. Visualizations

Figure 4. Check-boxes used to choose anchoring concepts for analyses on the EDAS. Kinetics and equilibrium anchoring concepts are not found on the GC12F exam, and therefore, cannot be chosen.

EDAS Requirements and Structure

Figure 6 shows the specific data formats that are required for using the EDAS. Student data must be in a *wide format*; this means that students are represented by rows and columns represent their responses to each of the 70 questions. Because the EDAS conducts several item level analyses, instructors must copy-and-paste students' A, B, C, and D responses per item into the spreadsheet. As an example, Figure 6 shows a segment of the spreadsheet where the data is inputted. Highlighted in yellow are the students' responses to the individual questions.

You have chosen:

70

out of

70

questions

Select All

Unselect All

Choose your Questions:

Questions (Gray Form)				Questions (Yellow Form)			
<input checked="" type="checkbox"/> Q1	<input checked="" type="checkbox"/> Q19	<input checked="" type="checkbox"/> Q37	<input checked="" type="checkbox"/> Q55	<input checked="" type="checkbox"/> Q1	<input checked="" type="checkbox"/> Q19	<input checked="" type="checkbox"/> Q37	<input checked="" type="checkbox"/> Q55
<input checked="" type="checkbox"/> Q2	<input checked="" type="checkbox"/> Q20	<input checked="" type="checkbox"/> Q38	<input checked="" type="checkbox"/> Q56	<input checked="" type="checkbox"/> Q2	<input checked="" type="checkbox"/> Q20	<input checked="" type="checkbox"/> Q38	<input checked="" type="checkbox"/> Q56
<input checked="" type="checkbox"/> Q3	<input checked="" type="checkbox"/> Q21	<input checked="" type="checkbox"/> Q39	<input checked="" type="checkbox"/> Q57	<input checked="" type="checkbox"/> Q3	<input checked="" type="checkbox"/> Q21	<input checked="" type="checkbox"/> Q39	<input checked="" type="checkbox"/> Q57
<input checked="" type="checkbox"/> Q4	<input checked="" type="checkbox"/> Q22	<input checked="" type="checkbox"/> Q40	<input checked="" type="checkbox"/> Q58	<input checked="" type="checkbox"/> Q4	<input checked="" type="checkbox"/> Q22	<input checked="" type="checkbox"/> Q40	<input checked="" type="checkbox"/> Q58
<input checked="" type="checkbox"/> Q5	<input checked="" type="checkbox"/> Q23	<input checked="" type="checkbox"/> Q41	<input checked="" type="checkbox"/> Q59	<input checked="" type="checkbox"/> Q5	<input checked="" type="checkbox"/> Q23	<input checked="" type="checkbox"/> Q41	<input checked="" type="checkbox"/> Q59
<input checked="" type="checkbox"/> Q6	<input checked="" type="checkbox"/> Q24	<input checked="" type="checkbox"/> Q42	<input checked="" type="checkbox"/> Q60	<input checked="" type="checkbox"/> Q6	<input checked="" type="checkbox"/> Q24	<input checked="" type="checkbox"/> Q42	<input checked="" type="checkbox"/> Q60
<input checked="" type="checkbox"/> Q7	<input checked="" type="checkbox"/> Q25	<input checked="" type="checkbox"/> Q43	<input checked="" type="checkbox"/> Q61	<input checked="" type="checkbox"/> Q7	<input checked="" type="checkbox"/> Q25	<input checked="" type="checkbox"/> Q43	<input checked="" type="checkbox"/> Q61
<input checked="" type="checkbox"/> Q8	<input checked="" type="checkbox"/> Q26	<input checked="" type="checkbox"/> Q44	<input checked="" type="checkbox"/> Q62	<input checked="" type="checkbox"/> Q8	<input checked="" type="checkbox"/> Q26	<input checked="" type="checkbox"/> Q44	<input checked="" type="checkbox"/> Q62
<input checked="" type="checkbox"/> Q9	<input checked="" type="checkbox"/> Q27	<input checked="" type="checkbox"/> Q45	<input checked="" type="checkbox"/> Q63	<input checked="" type="checkbox"/> Q9	<input checked="" type="checkbox"/> Q27	<input checked="" type="checkbox"/> Q45	<input checked="" type="checkbox"/> Q63
<input checked="" type="checkbox"/> Q10	<input checked="" type="checkbox"/> Q28	<input checked="" type="checkbox"/> Q46	<input checked="" type="checkbox"/> Q64	<input checked="" type="checkbox"/> Q10	<input checked="" type="checkbox"/> Q28	<input checked="" type="checkbox"/> Q46	<input checked="" type="checkbox"/> Q64
<input checked="" type="checkbox"/> Q11	<input checked="" type="checkbox"/> Q29	<input checked="" type="checkbox"/> Q47	<input checked="" type="checkbox"/> Q65	<input checked="" type="checkbox"/> Q11	<input checked="" type="checkbox"/> Q29	<input checked="" type="checkbox"/> Q47	<input checked="" type="checkbox"/> Q65
<input checked="" type="checkbox"/> Q12	<input checked="" type="checkbox"/> Q30	<input checked="" type="checkbox"/> Q48	<input checked="" type="checkbox"/> Q66	<input checked="" type="checkbox"/> Q12	<input checked="" type="checkbox"/> Q30	<input checked="" type="checkbox"/> Q48	<input checked="" type="checkbox"/> Q66
<input checked="" type="checkbox"/> Q13	<input checked="" type="checkbox"/> Q31	<input checked="" type="checkbox"/> Q49	<input checked="" type="checkbox"/> Q67	<input checked="" type="checkbox"/> Q13	<input checked="" type="checkbox"/> Q31	<input checked="" type="checkbox"/> Q49	<input checked="" type="checkbox"/> Q67
<input checked="" type="checkbox"/> Q14	<input checked="" type="checkbox"/> Q32	<input checked="" type="checkbox"/> Q50	<input checked="" type="checkbox"/> Q68	<input checked="" type="checkbox"/> Q14	<input checked="" type="checkbox"/> Q32	<input checked="" type="checkbox"/> Q50	<input checked="" type="checkbox"/> Q68
<input checked="" type="checkbox"/> Q15	<input checked="" type="checkbox"/> Q33	<input checked="" type="checkbox"/> Q51	<input checked="" type="checkbox"/> Q69	<input checked="" type="checkbox"/> Q15	<input checked="" type="checkbox"/> Q33	<input checked="" type="checkbox"/> Q51	<input checked="" type="checkbox"/> Q69
<input checked="" type="checkbox"/> Q16	<input checked="" type="checkbox"/> Q34	<input checked="" type="checkbox"/> Q52	<input checked="" type="checkbox"/> Q70	<input checked="" type="checkbox"/> Q16	<input checked="" type="checkbox"/> Q34	<input checked="" type="checkbox"/> Q52	<input checked="" type="checkbox"/> Q70
<input checked="" type="checkbox"/> Q17	<input checked="" type="checkbox"/> Q35	<input checked="" type="checkbox"/> Q53		<input checked="" type="checkbox"/> Q17	<input checked="" type="checkbox"/> Q35	<input checked="" type="checkbox"/> Q53	
<input checked="" type="checkbox"/> Q18	<input checked="" type="checkbox"/> Q36	<input checked="" type="checkbox"/> Q54		<input checked="" type="checkbox"/> Q18	<input checked="" type="checkbox"/> Q36	<input checked="" type="checkbox"/> Q54	

Figure 3. Check-boxes used to choose subsets of questions for the EDAS analyses.

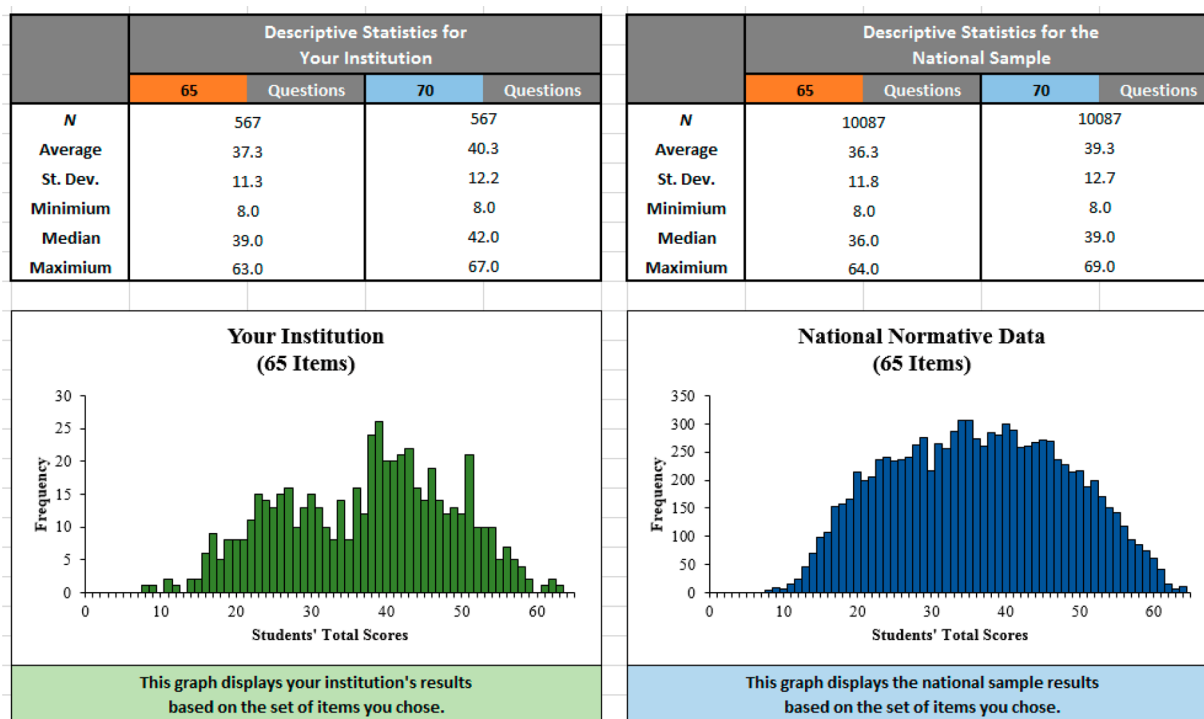


Figure 5. Example of some of the results provided when using the EDAS.

ID	Form	Q1	Q2	Q3	Q4	Q5
KEY	Y	*	*	*	*	*
KEY	G	*	*	*	*	*
1	G	A	B	C	D	A
2	G	B	C	D	A	B
3	G	C	D	A	B	C
4	G	D	A	B	C	D
5	G	A	B	C	D	A
6	G	B	C	D	A	B
7	G	C	D	A	B	C
8	G	D	A	B	C	D
9	G	A	B	C	D	A
10	Y	B	C	D	A	B
11	Y	C	D	A	B	C
12	Y	D	A	B	C	D
13	Y	A	B	C	D	A
14	Y	B	C	D	A	B
15	Y	C	D	A	B	C

Figure 6. Example of how data needs to be formatted for the EDAS. The answers were replaced with * to preserve the security of the exam answer key.

The answer keys for the gray (G) and yellow (Y) versions of the exam must be inputted using capital letters in the two rows labeled "Key Y" and "Key G". The answers themselves have been omitted from Figure 6 in order to preserve the security of the exams. For exams that have two separate forms, the user must also enter whether the students took the G or Y versions, as shown in blue in Figure 6. The EDAS can only process up to 3,000 students. For data sets larger than this, it is recommended that a subset or a random sample of students is used for analyses.

The EDAS calculations will only recognize item responses and answer keys that are entered using capital letters. If a cell is

left blank or includes any other values, the EDAS will consider this either an incorrect response or a nonresponse (*missing data*).²⁰ The national normative data sets assembled by the ACS-EI are collected from volunteering professors from around the country. As a result, it is not uncommon for instructors to send students' total scores, but not their responses to individual items. For this reason, the ACS-EI has used methods to estimate the missing item response data used for the national norm statistics.²⁰

The EDAS performs best when it is used with recent versions of MS Excel,¹⁷ which includes MS Excel 2010 and 2013 (PC) or MS Excel 2011 (Mac). Some features may lose their functionality if the EDAS is used with other versions. The features within the EDAS are divided across multiple spreadsheets, and instructors can navigate through the different features using clickable buttons or spreadsheet tabs. An instruction manual was developed in order to help instructors navigate through the EDAS features.

Security

The EDAS includes secure and confidential information about ACS exams and national normative results. Therefore, instructors wishing to obtain a copy of the EDAS must take all necessary steps to maintain this security and confidentiality including storing EDAS in a secured location and on a password protected computer. The steps necessary to maintain security are similar to those necessary to protect individual ACS exams, and as a result, instructors must acknowledge and sign a nondisclosure security agreement in order to obtain a copy of the EDAS. For additional security, the EDAS must be delivered via a method that allows for a signature.

EDAS TRIAL TESTING

In order to incorporate faculty feedback into the EDAS, trial testing was conducted to help improve the usability of the interface and to evaluate its usefulness as an assessment analysis

tool. A total of 10 chemistry instructors were recruited to participate; these instructors were either participants in previous ACS-EI workshops or they volunteered while at a recent ACS National Meeting and Exposition. For the purposes of trial testing, the instructors used their own computers, and they reported using either MS Excel versions 2010 (5 participants), 2011 (1 participant), or 2013 (4 participants). The volunteering participants were chemistry faculty who taught in a variety of different chemistry departments including 2 A.S. (Associate of Science degree) granting departments, 6 B.S. and B.A. granting departments, and 2 Ph.D. granting departments. Additionally, the instructors taught at colleges/universities from multiple geographic locations across the U.S., including one from the West, three from the Midwest, one from the Northeast, and five from the South. IRB approval was granted prior to collecting survey data and all of the instructors consented for their data to be used for trial testing purposes.

The following research is considered market research, and therefore, the ultimate goal was to develop a product that was useful to ACS exam users, rather than to generalize the results to a larger population. As a result, feedback was collected on a small-scale in order to receive more personalized information from a few participants. The EDAS was copied onto a CD or USB flash drive and was sent to the participants via mail. The participants were sent an EDAS created for the GC12F exam; however, not all of the participants used this exam. Therefore, a mock student data set was sent with the EDAS, and the instructors were asked to think about this as a prototype for a version of an ACS exam that they specifically used. In order to elicit feedback, an online survey was emailed to the participants using Qualtrics.²¹ The survey began with the System Usability Scale (SUS), which is a ten-item survey that uses a five-point Likert scale to elicit user feedback about the usability of a system.²² For this study, a modified version of the SUS was used that replaced the term “system” with “product.”^{23–26} The validity and reliability of the data produced by the original and modified versions of the SUS have been extensively studied in the literature with numerous different samples and products.^{22–31}

The SUS can be used to determine how users perceive the usability of a system or product, but it is not meant to be used to diagnose usability problems.²⁸ Therefore, the SUS was given as a way to initially prime instructors to think about the usability of the EDAS and additional items were developed to elicit feedback specific to the individual EDAS features. A majority of the diagnostic items were either open-ended or required the instructors to choose multiple choices from a list of options. The full survey took instructors approximately 15–20 min to complete. Each participant had an open-line of email communication with the researchers, and the participants were urged to send questions, additional comments, or concerns about the EDAS, the associated instruction manual, or the survey via email. None of the instructors expressed confusion regarding any of the survey items. Additionally, consistency was noted across the individual instructor's responses to the SUS and the diagnostic items; this suggests that the instructors' feedback was collected in a robust manner, even though the sample size may be too small for formal psychometric evaluation of the data.

Utility and Usability

The overarching goal of the trial testing was to identify areas where the EDAS needed improvement in terms of its utility for

departmental/classroom assessment or the usability of its user interface. The general pattern of feedback suggested that instructors found the EDAS useful for their assessment purposes. This was evident by the instructors' responses to the SUS item that states *I think that I would like to use this product frequently* (shown in Figure 7a). However, the

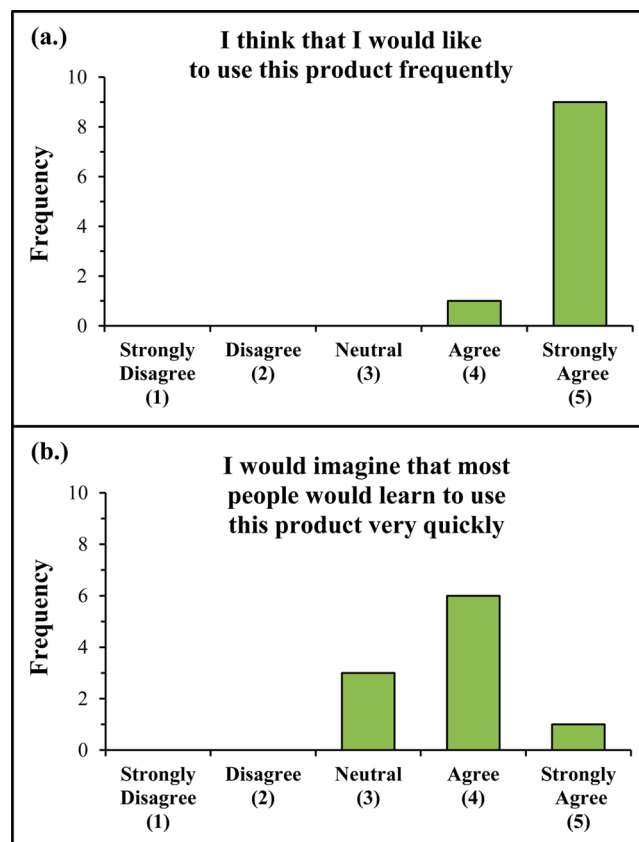


Figure 7. Instructors' responses to items 1 and 7 on the SUS.

instructors also described that there were some limitations in terms of the usability. Figure 7b shows the responses to the item that states *I would imagine that most people would learn to use this product very quickly*. Although none of the instructors disagreed with this statement, three reported feeling neutral. Graphical displays for the additional SUS items can be found in the Supporting Information.

Table 1 displays the instructors' raw responses to the SUS items. As a result, it would be expected that positively worded items would have responses near the high end of the scale (4 or 5) and negatively worded items would have responses near the low end of the scale (1 or 2), and this trend is generally observed in Table 1. After recoding the five negatively worded items in the positive direction, the average instructor response across the items was a 4.1, which is arguably a relatively high score on the 1–5 Likert scale.

Although the SUS scores provided a broad overview of the instructors' perceptions about the EDAS utility and usability, additional items focused more specifically on the individual EDAS features. Table 2 shows the results of two such items (shown in Figures 8 and 9). These items asked instructors to indicate which EDAS features they found helpful. Figure 9 also had an analogous item that was related to choosing sets of individual questions (as shown in Figure 3) rather than

Table 1. Reported Usability of the EDAS As Provided by the SUS^a

Item Number ^b	Item Statement	Average	SD	Min	Median	Max
1(p)	I think that I would like to use this product frequently.	4.9	0.32	4.0	5.0	5.0
3(p)	I thought the product was easy to use.	4.0	0.82	3.0	4.0	5.0
5(p)	I found that the various functions in this product were well integrated.	4.1	0.57	3.0	4.0	5.0
7(p)	I would imagine that most people would learn to use this product very quickly.	3.8	0.63	3.0	4.0	5.0
9(p)	I felt very confident using the product.	4.2	0.92	2.0	4.0	5.0
2(n)	I found the product unnecessarily complex.	2.2	0.92	1.0	2.0	4.0
4(n)	I think that I would need the support of a technical person to be able to use this product.	1.4	0.52	1.0	1.0	2.0
6(n)	I thought that there was too much inconsistency in this product.	1.7	0.67	1.0	2.0	3.0
8(n)	I found the product very awkward to use.	2.2	0.92	1.0	2.0	4.0
10(n)	I needed to learn a lot of things before I could get going with this product.	2.3	0.95	1.0	2.0	4.0

^a*n* = 10. ^b(p) represents a positively worded item; (n) represents a negatively worded item.

Table 2. Frequency of Instructors Who Found the EDAS Features Helpful

EDAS Features	Number of Participants (<i>n</i> = 10)
Descriptive Statistics	10
Percentiles	10
Difficulty	10
Discrimination	8
Individual Item Bar Plots	10
Select Individual Items	8
Select Anchoring Concepts (ACCM)	8

Which statistical analyses featured provided by the EDAS would be helpful for you? (check all that apply)

☐ Descriptive Statistics (e.g., means, standard deviations, etc.)

☐ Percentiles

☐ Difficulty

☐ Discrimination

☐ Individual Item Responses

☐ None of the features would be helpful

☐ Other

Figure 8. Trial testing survey item about helpful EDAS features.

One key feature of the EDAS is to be able to select anchoring concepts for data analysis. Is it helpful for you to be able to see the analyses for anchoring concepts? Please briefly explain your response.

☐ Yes

☐ No

Figure 9. Trial testing survey item about selecting anchoring concepts.

anchoring concepts (Figure 4). A total of 8 out of 10 professors reported that all of the statistical features (Figure 8) on the

EDAS were helpful, and two instructors indicated that discrimination would not be useful to them. Similarly, 8 out of 10 instructors indicated that selecting subsets of items and anchoring concepts would be useful to them.

The ability to select subsets of items or anchoring concepts for analyses are two highly novel features of the EDAS, and in most cases, the instructors found these aspects helpful for their assessment purposes. In regards to the anchoring concepts one instructor described, “*This can help the instructor make decisions about what to emphasize/deemphasize in the course to meet the course goals.*” In terms of selecting individual items, instructors viewed two ways to use this feature: (1) they could examine a subsets of items based on a given topic of interest or (2) they could remove questions from analyses when the course did not cover that content. As an example an instructor described, “*I frequently like to analyze my student performance on given topics, particularly when only a subset of the questions was actually covered in the course at the given time. I could also select on my own questions that I thought best evaluated student comprehension on a given topic to compare my students to national norms.*” Further, instructors suggested additional features that they believed would be useful for the EDAS including long-term trend analyses, analyses of course sections, and reporting the analyses of all of the big ideas at once. At this time, these additions would substantially increase the file size of the EDAS, which may begin to make it difficult to use on some computers. As a result, these aspects will be considered for future versions of the EDAS or when the functionality is ported to a different application that does not use MS Excel.

The instructors also responded with feedback for how to improve the usability of the EDAS. As a few examples, some instructors spoke to difficulty navigating the EDAS interface or using MS Excel, whereas others spoke to the need for enhanced explanations for difficulty and discrimination indices. However, most of the reported challenges were focused on entering data into the EDAS. Some of the difficulties included using the *paste values* option in MS Excel, finding the right columns to paste the values into, and inputting the “G” and “Y” to indicate which version of the exam was used. While some of these challenges resulted in revisions to the EDAS and the user interface, others were much more difficult to resolve. As a result, a video user guide was developed that should explicitly provide explanations to the difficulties reported so that users may better visualize and perform these processes.

In order to assess the utility of the EDAS as an assessment tool, instructors were asked to respond to the open-ended question Do you think that the EDAS could be helpful to you? If so, please explain how you may be able to use the results

provided by the EDAS? An overview of the responses given by instructors can be found in Table 3. A total of 9/10 participants

Table 3. Instructors' Responses to the Question: Do You Think That the EDAS Could Be Helpful to You? If so, Please Explain How You May Be Able to Use the Results Provided by the EDAS?

Responses	Number of Participants (n = 10) ^a
EDAS is useful	9
Value making national comparisons	6
Help identify course improvements or evaluate whether students have met learning outcomes	5
Appreciate selecting specific items	3
Appreciate selecting anchoring concepts	3
EDAS is quick, efficient, and/or easy to use	4
Help organize results	1

^aParticipants may have discussed more than one response category.

described that the EDAS would be useful for them, and these instructors often spoke toward the potential to make comparisons to the national sample, to make classroom improvements, or to assess course learning goals. Most of the instructors expressed their excitement to use the EDAS tool, and in general, many viewed the EDAS as having strong implications. As an example two instructors described:

On the whole, I think the EDAS is a great tool. I believe that most instructors would be very keen to know how their students are doing compared to the national average. This may also help individual instructors to try and improve certain aspects of their own teaching methodologies which, in the long term, obviously benefits the general student population.

Being able to make connections on the anchoring concepts between how my students do compared to the national database is invaluable! Also, the flexibility to select individual questions to compare or specific anchoring concepts allows for more in-depth analysis.

CONCLUSIONS

In the recent years, departmental assessment efforts have become more common than ever.^{4,5} Despite this commonality, literature has reported challenges for instructors when it comes to assessment data and using the results of those efforts to improve instruction.^{32,33} For example, chemistry departments interested in receiving ACS program approval will need to consider the assessment recommendations inherent in the 2015 guidelines reported by the Committee on Professional Training.^{15,16} ACS exams are a reliable way to assess such guidelines, however, efforts that focus on overall test scores may not be sensitive enough to evaluate specific content goals. As a result, the EDAS is a tool that instructors can use to help explore student performances based on subsets of exam questions or based on the anchoring concepts from the ACCM.^{10–14} The EDAS also allows instructors to compare their student performances to that of the national norms. Because the ACS-EI collects national normative data sets on a volunteer basis, professors sometimes send students' overall test level performances but not their responses on an individual item basis. As a result, methods to estimate the missing data were used by the ACS-EI,²⁰ and the EDAS uses these

augmented data sets to provide the national normative comparisons.

In order to assess its utility and usability, the EDAS went through trial testing with 10 chemistry instructors from across the country and most of the instructors stated that the EDAS was a useful tool with several implications for the classroom. However, some difficulties were noted in terms of the usability of the EDAS interface. As a result, several revisions were made, and a video user-guide was developed that has been tailored to the specific challenges described by the instructors. At this time, a version of the EDAS is developed for three ACS exams including the 2012 First Term General Chemistry Exam (GC12F), the 2010 Second Term General Chemistry Exam (GC10S), and the 2013 (Full Year) General Chemistry Exam (GC13). Instructors interested in obtaining a copy of the EDAS should contact The Institute at chmexams@iastate.edu.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available on the ACS Publications website at DOI: [10.1021/acs.jchemed.5b00474](https://doi.org/10.1021/acs.jchemed.5b00474).

The Supporting Information includes the instructors' responses to the eight additional items on the SUS that were not reported in the body of the manuscript. (PDF, DOCX)

AUTHOR INFORMATION

Corresponding Author

*E-mail: taholme@iastate.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We would like to sincerely thank the instructors who participated in the EDAS trial testing for their time and incredibly valuable feedback! This work was supported by NSF DUE #1323288; any opinions, findings, conclusions and recommendations expressed in this report are those of the authors and do not necessarily reflect the views of the National Science Foundation (NSF).

REFERENCES

- (1) Pienta, N. J. Striking a Balance with Assessment. *J. Chem. Educ.* **2011**, *88* (9), 1199–1200.
- (2) Towns, M. H. Developing Learning Objectives and Assessment Plans at a Variety of Institutions: Examples and Case Studies. *J. Chem. Educ.* **2010**, *87* (1), 91–96.
- (3) Bretz, S. L. Navigating the Landscape of Assessment. *J. Chem. Educ.* **2012**, *89* (6), 689–691.
- (4) Emenike, M. E.; Schroeder, J. D.; Murphy, K. L.; Holme, T. A Snapshot of Chemistry Faculty Members' Awareness of Departmental Assessment Efforts. *Assess. Update*. **2011**, *23* (4), 1–2 14–16..
- (5) Emenike, M. E.; Schroeder, J.; Murphy, K.; Holme, T. Results from a National Needs Assessment Survey: A View of Assessment Efforts within Chemistry Departments. *J. Chem. Educ.* **2013**, *90* (5), 561–567.
- (6) Emenike, M.; Raker, J. R.; Holme, T. Validating Chemistry Faculty Members' Self-Reported Familiarity with Assessment Terminology. *J. Chem. Educ.* **2013**, *90* (9), 1130–1136.
- (7) Raker, J. R.; Emenike, M. E.; Holme, T. A. Using Structural Equation Modeling to Understand Chemistry Faculty Familiarity of Assessment Terminology: Results from a National Survey. *J. Chem. Educ.* **2013**, *90* (8), 981–987.

- (8) Raker, J. R.; Holme, T. A. Investigating Faculty Familiarity with Assessment Terminology by Applying Cluster Analysis to Interpret Survey Data. *J. Chem. Educ.* **2014**, *91* (8), 1145–1151.
- (9) National Research Council. *Discipline-Based Education Research: Understanding and Improving Learning in Undergraduate Science and Engineering*; Singer, S. R., Nielsen, N. R., Schweingruber, H. A., Eds.; The National Academies Press: Washington, DC, 2012.
- (10) Holme, T.; Murphy, K. The ACS Exams Institute Undergraduate Chemistry Anchoring Concepts Content Map I: General Chemistry. *J. Chem. Educ.* **2012**, *89* (6), 721–723.
- (11) Raker, J.; Holme, T.; Murphy, K. The ACS Exams Institute Undergraduate Chemistry Anchoring Concepts Content Map II: Organic Chemistry. *J. Chem. Educ.* **2013**, *90* (11), 1443–1445.
- (12) Holme, T.; Luxford, C.; Murphy, K. Updating the General Chemistry Anchoring Concepts Content Map. *J. Chem. Educ.* **2015**, *92* (6), 1115–1116.
- (13) Luxford, C. J.; Linenberger, K. J.; Raker, J. R.; Baluyut, J. Y.; Reed, J. J.; De Silva, C.; Holme, T. A. Building a Database for the Historical Analysis of the General Chemistry Curriculum Using ACS General Chemistry Exams as Artifacts. *J. Chem. Educ.* **2015**, *92* (2), 230–236.
- (14) Luxford, C. J.; Holme, T. A. What Do Conceptual Holes in Assessment Say about the Topics we Teach in General Chemistry? *J. Chem. Educ.* **2015**, *92* (6), 993–1002.
- (15) Undergraduate Professional Education in Chemistry: ACS Guidelines and Evaluation Procedures for Bachelor's Degree Programs. <http://www.acs.org/content/dam/acsorg/about/governance/committees/training/2015-acg-guidelines-for-bachelors-degree-programs.pdf> (accessed June, 2015).
- (16) Wenzel, T. J.; McCoy, A. B.; Landis, C. R. An Overview of the Changes in the 2015 ACS Guidelines for Bachelor's Degree Programs. *J. Chem. Educ.* **2015**, *92* (6), 965–968.
- (17) Microsoft Excel, Computer Software; Microsoft, Inc.: Redmond, WA, 2010/2011/2013.
- (18) Walkenbach, J. *Microsoft Excel 2013: Power Programming with VBA*; John Wiley & Sons, Inc.: Hoboken, NJ, 2013.
- (19) SAS Institute Inc. SAS 9.4 Software. http://www.sas.com/en_us/ (accessed June, 2015).
- (20) Brandriet, A.; Holme, T. Methods for Addressing Missing Data with Applications from ACS Exams. *J. Chem. Educ.* **2015**, DOI: 10.1021/acs.jchemed.5b00180.
- (21) Qualtrics: *Online Survey Software*; Qualtrics: Provo, UT, 2015.
- (22) Brooke, J. SUS: A "Quick and Dirty" Usability Scale. In *Usability Evaluation in Industry*; Jordan, P. W., Thomas, B., Weerdmeester, B. A., McClelland, I. L., Eds.; Taylor & Francis: Bristol, PA, 1996; p 189–194.
- (23) Bangor, A.; Kortum, P. T.; Miller, J. T. An Empirical Evaluation of the System Usability Scale. *Int. J. Hum-Comput. Int.* **2008**, *24* (6), 574–594.
- (24) Bangor, A.; Kortum, P.; Miller, J. Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. *J. Usability Stud.* **2009**, *4* (3), 114–123.
- (25) Lewis, J. R.; Sauro, J. The Factor Structure of the System Usability Scale. In *Human Centered Design*, Proceedings of the First International Conference, HCD 2009, Held as Part of HCI International; Kurosu, M., Ed.; Springer: Berlin Heidelberg, 2009; pp 94–103.
- (26) Kortum, P. T.; Bangor, A. Usability Ratings for Everyday Products Measured with the System Usability Scale. *Int. J. Hum-Comput. Int.* **2013**, *29* (2), 67–76.
- (27) Borsci, S.; Federici, S.; Lauriola, M. On the Dimensionality of the System Usability Scale: A Test of Alternative Measurement Models. *Cogn. Process.* **2009**, *10* (3), 193–197.
- (28) Sauro, J. SUSified? Little-Known System Usability Scale Facts. *User Experience Magazine*, August 2011. <http://uxpamagazine.org/sustified/> (accessed June 2015).
- (29) Brooke, J. SUS: A Retrospective. *J. Usability Stud.* **2013**, *8* (2), 29–40.
- (30) Sauro, J. Measuring Usability with the System Usability Scale (SUS). <http://www.measuringu.com/sus.php> (accessed June, 2015).
- (31) U.S. Department of Health & Human Services. System Usability Scale (SUS). <http://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html> (accessed June, 2015).
- (32) Banta, T. W.; Blaich, C. Closing the Assessment Loop. *Change* **2010**, *43* (1), 22–27.
- (33) Hutchings, P. From Departmental to Disciplinary Assessment: Deepening Faculty Engagement. *Change* **2011**, *43* (5), 36–43.